



500.43123X00

Applicant(s): Koichi UENO

Serial No.: 10/658,395

Filed: September 10, 2003

Title: STORAGE DEVICE AND STORAGE DEVICE CONTROL METHOD

LETTER CLAIMING RIGHT OF PRIORITY

Commissioner for Patents
P.O. Box 1450
Alexandria, VA 22313-1450

September 25, 2003

Sir:

Under the provisions of 35 USC 119 and 37 CFR 1.55, the applicant(s) hereby
claim(s) the right of priority based on:

Japanese Patent Application No. 2003-185305
Filed: June 27, 2003

A certified copy of said Japanese Patent Application is attached.

Respectfully submitted,

ANTONELLI, TERRY, STOUT & KRAUS, LLP

Carl I. Brundidge
Registration No. 29,621

CIB/rp
Attachment

日 本 国 特 許 庁
JAPAN PATENT OFFICE

別紙添付の書類に記載されている事項は下記の出願書類に記載されている事項と同一であることを証明する。

This is to certify that the annexed is a true copy of the following application as filed with this Office.

出 願 年 月 日 2 0 0 3 年 6 月 2 7 日
Date of Application:

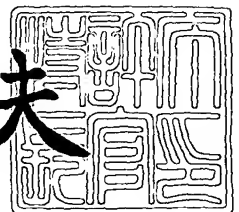
出 願 番 号 特 願 2 0 0 3 - 1 8 5 3 0 5
Application Number:
[ST. 10/C]: [J P 2 0 0 3 - 1 8 5 3 0 5]

出 願 人 株式会社日立製作所
Applicant(s):

2 0 0 3 年 8 月 2 6 日

特許庁長官
Commissioner,
Japan Patent Office

今 井 康 夫



出証番号 出証特 2 0 0 3 - 3 0 6 9 6 4 8

【書類名】 特許願

【整理番号】 HI030166

【提出日】 平成15年 6月27日

【あて先】 特許庁長官殿

【国際特許分類】 G06F 3/06

【発明者】

【住所又は居所】 神奈川県小田原市中里 3 2 2 番 2 号 株式会社日立製作所 R A I D システム事業部内

【氏名】 上野 浩一

【特許出願人】

【識別番号】 000005108

【氏名又は名称】 株式会社日立製作所

【代理人】

【識別番号】 110000176

【氏名又は名称】 一色国際特許業務法人

【代表者】 一色 健輔

【手数料の表示】

【予納台帳番号】 211868

【納付金額】 21,000円

【提出物件の目録】

【物件名】 明細書 1

【物件名】 図面 1

【物件名】 要約書 1

【プルーフの要否】 要

【書類名】 明細書

【発明の名称】 ストレージ装置、及びストレージ装置の制御方法

【特許請求の範囲】

【請求項 1】 データを記憶するための複数の記憶ボリュームと、ある時点の後に、第 1 の前記記憶ボリュームに記憶される前記ある時点における前記データの更新要求を受信する手段と、前記更新要求の対象となった前記ある時点における前記データの複製を第 2 の前記記憶ボリュームに記憶する手段とを備え、

前記第 2 の記憶ボリュームの状態に応じて、前記更新要求の対象となった前記ある時点における前記データの複製を、第 3 の前記記憶ボリュームの未使用の記憶領域に記憶する手段

を備えることを特徴とするストレージ装置。

【請求項 2】 請求項 1 に記載のストレージ装置において、

前記第 2 の記憶ボリュームの状態に応じて、前記更新要求の対象となった前記ある時点における前記データの複製を、前記第 3 の記憶ボリュームの未使用の記憶領域に記憶する前記手段は、

前記第 2 の記憶ボリュームの未使用の記憶領域の記憶容量が判定値よりも小さい場合に、前記更新要求の対象となった前記ある時点における前記データの複製を、前記第 3 の記憶ボリュームの未使用の記憶領域に記憶する手段であることを特徴とするストレージ装置。

【請求項 3】 請求項 2 に記載のストレージ装置において、

前記第 2 の記憶ボリュームの未使用の記憶領域の記憶容量をユーザインタフェースに表示する手段

を備えることを特徴とするストレージ装置。

【請求項 4】 請求項 2 に記載のストレージ装置において、

前記第 2 の記憶ボリュームの未使用の記憶領域の記憶容量に応じて、ユーザインタフェースに警告を表示する手段

を備えることを特徴とするストレージ装置。

【請求項 5】 請求項 1 に記載のストレージ装置において、

ユーザインタフェースから入力された前記第 3 の記憶ボリュームを指定するた

めの情報に応じて、前記第3の記憶ボリュームを選定する手段を備えることを特徴とするストレージ装置。

【請求項6】 請求項1に記載のストレージ装置において、

前記ストレージ装置が備える各前記記憶ボリュームの未使用の記憶領域の記憶容量に応じて、前記第3の記憶ボリュームを選定する手段を備えることを特徴とするストレージ装置。

【請求項7】 請求項1に記載のストレージ装置において、

前記更新要求の対象となった前記ある時点における前記データの複製を前記第2の記憶ボリュームに記憶する前記手段は、

前記第2の記憶ボリュームの記憶領域のうち前記データの複製が記憶される記憶領域の部分を、前記データを記憶するための区画に区分する手段と、

前記区画に区分された前記記憶領域の部分に前記データの複製を書き込む手段と

であることを特徴とするストレージ装置。

【請求項8】 請求項1に記載のストレージ装置において、

前記第2の記憶ボリュームの状態に応じて、前記更新要求の対象となった前記ある時点における前記データの複製を、前記第3の記憶ボリュームの未使用の記憶領域に記憶する前記手段は、

前記第2の記憶ボリュームの状態に応じて、前記更新要求の対象となった前記ある時点における前記データの複製を、前記第3の記憶ボリュームの前記未使用の記憶領域うち前記データの複製が記憶される前記未使用の記憶領域の部分を、前記データを記憶するための区画に区分する手段と、

前記区画に区分された前記未使用の記憶領域の部分に前記データの複製を書き込む手段と

であることを特徴とするストレージ装置。

【請求項9】 データを記憶するための複数の記憶ボリュームを備えるストレージ装置の制御方法であって、

ある時点の後に、第1の前記記憶ボリュームに記憶される前記ある時点における前記データの更新要求を受信するステップと、

前記更新要求の対象となった前記ある時点における前記データの複製を第2の前記記憶ボリュームに記憶するステップと、

前記第2の記憶ボリュームの状態に応じて、前記更新要求の対象となった前記ある時点における前記データの複製を、第3の前記記憶ボリュームの未使用の記憶領域に記憶するステップと

を備えることを特徴とするストレージ装置の制御方法。

【請求項10】 請求項9に記載のストレージ装置の制御方法において、

前記第2の記憶ボリュームの状態に応じて、前記更新要求の対象となった前記ある時点における前記データの複製を、前記第3の記憶ボリュームの未使用の記憶領域に記憶する前記ステップは、

前記第2の記憶ボリュームの未使用の記憶領域の記憶容量が判定値よりも小さい場合に、前記更新要求の対象となった前記ある時点における前記データの複製を、前記第3の記憶ボリュームの未使用の記憶領域に記憶するステップであること

を特徴とするストレージ装置の制御方法。

【請求項11】 請求項10に記載のストレージ装置の制御方法において、

前記第2の記憶ボリュームの未使用の記憶領域の記憶容量をユーザインタフェースに表示するステップ

を備えることを特徴とするストレージ装置の制御方法。

【請求項12】 請求項10に記載のストレージ装置の制御方法において、

前記第2の記憶ボリュームの未使用の記憶領域の記憶容量に応じて、ユーザインタフェースに警告を表示するステップ

を備えることを特徴とするストレージ装置の制御方法。

【請求項13】 請求項9に記載のストレージ装置の制御方法において、

ユーザインタフェースから入力された前記第3の記憶ボリュームを指定するための情報に応じて、前記第3の記憶ボリュームを選定するステップ

を備えることを特徴とするストレージ装置の制御方法。

【請求項14】 請求項9に記載のストレージ装置の制御方法において、

前記ストレージ装置が備える各前記記憶ボリュームの未使用の記憶領域の記憶

容量に応じて、前記第 3 の記憶ボリュームを選定するステップ
を備えることを特徴とするストレージ装置の制御方法。

【請求項 1 5】 請求項 9 に記載のストレージ装置の制御方法において、
前記更新要求の対象となった前記ある時点における前記データの複製を前記第
2 の記憶ボリュームに記憶する前記ステップは、
前記第 2 の記憶ボリュームの記憶領域のうち前記データの複製が記憶される記
憶領域の部分を、前記データを記憶するための区画に区分するステップと、
前記区画に区分された前記記憶領域の部分に前記データの複製を書き込むステ
ップと
であることを特徴とするストレージ装置の制御方法。

【請求項 1 6】 請求項 9 に記載のストレージ装置の制御方法において、
前記第 2 の記憶ボリュームの状態に応じて、前記更新要求の対象となった前記
ある時点における前記データの複製を、前記第 3 の記憶ボリュームの未使用の記
憶領域に記憶する前記ステップは、
前記第 2 の記憶ボリュームの状態に応じて、前記更新要求の対象となった前記
ある時点における前記データの複製を、前記第 3 の記憶ボリュームの前記未使用
の記憶領域うち前記データの複製が記憶される前記未使用の記憶領域の部分を、
前記データを記憶するための区画に区分するステップと、
前記区画に区分された前記未使用の記憶領域の部分に前記データの複製を書き
込むステップと
であることを特徴とするストレージ装置の制御方法。

【発明の詳細な説明】

【0 0 0 1】

【発明の属する技術分野】

本発明は、ストレージ装置、及びストレージ装置の制御方法に関する。

【0 0 0 2】

【従来の技術】

コンピュータシステムにおいては、ストレージ装置に記憶されるデータをバッ
クアップする場合など、元のデータを記憶した記憶ボリュームの複製を生成する

場合がある。この場合、元のデータを記憶した記憶ボリュームと複製データを記憶した記憶ボリュームとは内容が一致し整合性が保たれる必要がある。そのためには複製の生成が完了するまでの間、元のデータを記憶した記憶ボリュームの更新を停止させる必要がある。しかし高可用性が求められるコンピュータシステムなどでは、元のデータを記憶した記憶ボリュームに対するアクセスを停止することができない場合がある。

【0003】

そこで、元のデータを記憶した記憶ボリュームと複製データを記憶した記憶ボリュームとの間である時点での整合性を保ち、ある時点の後に元のデータが更新されたとしてもある時点の元のデータを参照できるようにすることができるスナップショットと呼ばれる技術が開発されている。

【0004】

スナップショットの技術によれば、整合性を保つべき時点の後に元のデータが更新される場合には、整合性を保つべき時点のデータを別の記憶ボリュームに記憶するようにする。つまり元のデータは、更新されなければそのままであり、更新される場合には整合性を保つべき時点のデータが別の記憶ボリュームに記憶される。

【0005】

【特許文献1】

特開 2001-306407 号公報

【0006】

【発明が解決しようとする課題】

しかしながら、通常、上記別の記憶ボリュームとして用意される記憶容量は、コスト抑制等のため元のデータが記憶される記憶ボリュームの記憶容量よりも小さい。

そのため従来のスナップショットの制御においては、元のデータを記憶した記憶ボリュームと複製データを記憶した記憶ボリュームとの整合性を保つことができる期間は、元のデータの更新頻度と上記別の記憶ボリュームの記憶容量と依存していた。

【0007】

本発明は上記課題を鑑みてなされたものであり、ストレージ装置、及びストレージ装置の制御方法を提供することを主たる目的とする。

【0008】**【課題を解決するための手段】**

上記課題を解決するために、本発明に係るストレージ装置は、データを記憶するための複数の記憶ボリュームと、ある時点の後に、第1の前記記憶ボリュームに記憶される前記ある時点における前記データの更新要求を受信する手段と、前記更新要求の対象となった前記ある時点における前記データの複製を第2の前記記憶ボリュームに記憶する手段とを備え、前記第2の記憶ボリュームの状態に応じて、前記更新要求の対象となった前記ある時点における前記データの複製を、第3の前記記憶ボリュームの未使用の記憶領域に記憶する手段を備える。

【0009】

これにより、例えば第2の記憶ボリューム（別の記憶ボリューム）の未使用記憶領域の記憶容量が判定値よりも小さくなった場合でも、第3の記憶ボリュームに前記データの複製を記憶させることにより、元のデータを記憶した記憶ボリュームと複製データを記憶した記憶ボリュームとの整合性を保つことができるようになる。また第1の記憶ボリュームに記憶されるデータに対する更新頻度に依存せずに、元のデータを記憶した記憶ボリュームと複製データを記憶した記憶ボリュームとの整合性を維持することができるようになる。

【0010】

ここで、記憶ボリュームとは、ハードディスク装置や半導体記憶装置等により構成されるディスクドライブにより提供される物理的な記憶領域である物理ボリュームと、物理ボリューム上に論理的に設定される記憶領域である論理ボリュームとを含む記憶リソースである。またスペアディスクを記憶ボリュームとして使用することも可能である。

【0011】

その他、本願が開示する課題、及びその解決方法は、発明の実施の形態の欄、及び図面により明らかにされる。

【0012】

【発明の実施の形態】

以下、本発明の実施の形態について図面を用いて詳細に説明する。

===全体構成例===

まず、本実施の形態に係るストレージ装置200を含む情報処理システムの全体構成を示すブロック図を図1に示す。本実施の形態に係る情報処理システムは情報処理装置1(100)、情報処理装置2(100)、ストレージ装置200、管理用装置250を備えている。

【0013】

===情報処理装置===

情報処理装置1(100)、情報処理装置2(100)はCPU(Central Processing Unit)やメモリを備えたコンピュータである。情報処理装置100が備えるCPUにより各種プログラムが実行されることにより、様々な機能が実現される。情報処理装置100は例えば銀行の自動預金預け払いシステムや航空機の座席予約システム等における中枢コンピュータとして利用される。情報処理装置100はメインフレームコンピュータとすることもできるし、パーソナルコンピュータとすることもできる。

【0014】

また情報処理装置100はストレージ装置200と通信可能に接続されており、ストレージ装置200に対してデータ入出力要求を送信する。データ入出力要求としては例えばデータの読み出し要求や書き込み要求(更新要求)などである。これにより情報処理装置100はストレージ装置200に記憶されているデータの読み出しや書き込みを行う。

【0015】

また情報処理装置100はストレージ装置200に対してストレージ装置200を管理するための各種コマンドの送信も行うようにすることができる。ストレージ装置200を管理するためのコマンドとしては、例えば後述するスナップショットコマンド等である。なお、情報処理装置100は2台に限られず1台とすることもできるし、3台以上とすることもできる。

【0016】

===ストレージ装置===

ストレージ装置200は、制御装置210、プライマリボリューム（第1の記憶ボリューム）220、セカンダリボリューム230、差分ボリューム（第2の記憶ボリューム）240を備えている。またストレージ装置200は管理用装置250に通信可能に接続されている。

【0017】

===記憶ボリューム===

プライマリボリューム220、差分ボリューム240は、データを記憶するための記憶ボリュームである。記憶ボリュームとは、ストレージ装置200が備えるハードディスク装置や半導体記憶装置等により構成されるディスクドライブにより提供される物理的な記憶領域である物理ボリュームと、物理ボリューム上に論理的に設定される記憶領域である論理ボリュームとを含む記憶リソースである。

【0018】

一方セカンダリボリューム230は、ストレージ装置200が仮想的に備える記憶ボリュームである。情報処理装置100はセカンダリボリューム230があたかも存在しているかのように、セカンダリボリューム230に対して例えばデータの読み出し要求を送信する。例えば図3に示すように、情報処理装置2（100）はセカンダリボリューム230から読み出したデータをテープ300等に記録しておくようにすることができる。

【0019】

===スナップショット制御===

この場合ストレージ装置200は、セカンダリボリューム230から読み出されるデータを、情報処理装置100から送信されるスナップショットコマンドに応じて制御することができる。すなわちストレージ装置200は、スナップショットコマンドに応じて、プライマリボリューム220に記憶されているデータと同一内容のデータがセカンダリボリューム230から読み出されるように制御することもできるし、ある時点のプライマリボリューム220に記憶されているデ

ータがセカンダリボリューム 230 から読み出されるように制御することもできる。この制御をスナップショット制御と呼ぶ。ある時点とは、例えばストレージ装置 200 がスナップショットコマンドを受信した時点とすることもできるし、スナップショットコマンドにより指示された時刻とすることもできる。

【0020】

前者の場合、すなわちセカンダリボリューム 230 から読み出されるデータを、プライマリボリューム 220 に記憶されているデータと同一内容のデータとする場合には、例えば情報処理装置 1 (100) がプライマリボリューム 220 に記憶されるデータを更新すると、ストレージ装置 200 は更新後のデータがセカンダリボリューム 230 にも反映されるように制御する。このときに情報処理装置 2 (100) からセカンダリボリューム 230 に記憶されているデータの読み出し要求を受信すると、ストレージ装置 200 はプライマリボリューム 220 に記憶されているデータを情報処理装置 2 (100) に送信する。これにより、情報処理装置 100 はセカンダリボリューム 230 からデータを読み出すことにより、プライマリボリューム 220 と同一のデータを取得することができる。

【0021】

一方後者の場合、すなわちセカンダリボリューム 230 から読み出されるデータを、ある時点のプライマリボリューム 220 に記憶されているデータとする場合には、例えば情報処理装置 1 (100) がプライマリボリューム 220 に記憶されるデータを更新すると、ストレージ装置 200 はプライマリボリューム 220 に記憶されていた更新前のある時点のデータを差分ボリューム 240 に記憶するように制御する。このときに情報処理装置 2 (100) からセカンダリボリューム 230 に記憶されている当該データに対する読み出し要求を受信すると、ストレージ装置 200 は差分ボリューム 240 に記憶されているデータを情報処理装置 2 (100) に送信する。これにより、情報処理装置 100 はセカンダリボリューム 230 からデータを読み出すことにより、プライマリボリューム 220 の更新有無に拘わらず、プライマリボリューム 220 のある時点のデータを取得することができる。なおこの場合、プライマリボリューム 220 に記憶される同一のデータが複数回更新されても、差分ボリューム 240 に記憶されるのは、あ

る時点のデータだけである。これにより、プライマリボリューム 220 には最新のデータが記憶され、差分ボリューム 240 にはある時点のデータが保持されることになる。またストレージ装置 200 が情報処理装置 1 (100) からプライマリボリューム 220 に記憶されるデータの更新要求を受信した場合に、更新後のデータを差分ボリューム 240 に記憶するようにすることもできる。この場合プライマリボリューム 220 には更新前のある時点のデータが記憶されることになる。

以上の処理の様子を示した図を図 4 に示す。

【0022】

なお上述の后者の場合には、プライマリボリューム 220 に記憶されるデータが更新されるにつれ、差分ボリューム 240 に記憶されるプライマリボリューム 220 のある時点のデータが増加する。そのため図 7 に示すように、差分ボリューム 240 の未使用記憶領域の記憶容量がなくなってしまった場合には、プライマリボリューム 220 に記憶されるある時点のデータの更新要求を受信しても、ストレージ装置 200 は、更新前のある時点のデータを差分ボリューム 240 に記憶することができない。

【0023】

本実施の形態においては、そのような場合には図 8 に示すように、差分ボリューム 240 とは別の他の記憶ボリューム (第 3 の記憶ボリューム) 241 を差分ボリューム 240 として使用し、その未使用記憶領域にプライマリボリューム 220 の更新前のある時点のデータを記憶するようにする。これにより、差分ボリューム 240 の未使用記憶領域の記憶容量が少なくなった場合であっても、ある時点のプライマリボリューム 220 のデータを保持し続けることができる。

【0024】

例えば図 3 に示したように、情報処理装置 2 (100) がセカンダリボリューム 240 に記憶されているデータを読み出してテープ 300 に書き込むことによりプライマリボリューム 220 のある時点のデータに対するバックアップ処理を行う場合に、ストレージ装置 200 は、差分ボリューム 240 の状態に応じて、プライマリボリューム 220 に記憶されているある時点のデータの複製を、差分

ボリューム 2 4 0 には記憶せずに、他の記憶ボリューム 2 4 1 の未使用記憶領域に記憶する。これにより、プライマリボリューム 2 2 0 のある時点のデータに対するバックアップ処理を継続することができるようになる。なお、差分ボリューム 2 4 0 の状態に応じてとは、差分ボリューム 2 4 0 の未使用記憶領域の記憶容量が判定値よりも小さくなった場合の他、未使用記憶領域の記憶容量の総記憶容量に対する割合が判定値よりも小さくなった場合や、使用済み記憶容量が判定値よりも大きくなった場合、あるいは使用済み記憶容量の総記憶容量に対する割合が判定値よりも大きくなった場合等も含む。

【 0 0 2 5 】

=== 制御装置 ===

制御装置 2 1 0 は本実施の形態に係る各種制御を実行するための手段として機能する。制御装置の構成を示すブロック図を図 2 に示す。

本実施の形態においては、制御装置 2 1 0 は、CPU (Central Processing Unit) 2 1 1、メモリ 2 1 2、ホスト I/F (InterFace) 2 1 3、ディスク I/F 2 1 4、内部 I/F 2 1 5 を備えている。

【 0 0 2 6 】

CPU 2 1 1 は制御装置 2 1 0 の全体の制御を司るものであり、メモリ 2 1 2 に記憶されたスナップショット制御プログラム 8 0 0 を実行する。メモリ 2 1 2 に記憶されたスナップショット制御プログラム 8 0 0 が CPU 2 1 1 により実行されることにより本実施の形態に係るスナップショット制御が行われる。

【 0 0 2 7 】

メモリ 2 1 2 にはスナップショット制御プログラム 8 0 0、スナップショット管理テーブル 4 0 0、S-VOL (Secondary VOLUME) 変換テーブル 5 0 0、ボリューム管理テーブル 6 0 0、差分ボリューム管理テーブル 7 0 0、フォーマット管理テーブル 9 0 0 が記憶されている。これらについては後述する。

【 0 0 2 8 】

ホスト I/F 2 1 3 は情報処理装置 1 0 0 と通信を行うための通信インタフェースとして機能する。情報処理装置 1 0 0 とストレージ装置 2 0 0 との間で行われる通信は、様々な通信プロトコルに従って行うようにすることができる。例え

ば、ファイバチャネルや S C S I (Small Computer System Interface)、F I C O N (Fibre Connection) (登録商標)、E S C O N (Enterprise System Connection) (登録商標)、A C O N A R C (Advanced Connection Architecture) (登録商標)、F I B A R C (Fibre Connection Architecture) (登録商標)、T C P / I P (Transmission Control Protocol/Internet Protocol) 等である。これらの通信プロトコルを混在させるようにすることもできる。例えば情報処理装置 1 (100) との間の通信はファイバチャネルで行い、情報処理装置 2 (100) との間の通信は T C P / I P で行うようにすることもできる。情報処理装置 100 がメインフレームコンピュータである場合には、例えば F I C O N や E S C O N、A C O N A R C、F I B A R C が用いられる。また情報処理装置 100 がオープン系の計算機である場合には、例えばファイバチャネルや S C S I、T C P / I P が用いられる。なお、情報処理装置 100 からのデータ入出力要求は、記憶ボリュームにおけるデータの管理単位であるブロックを単位として行うようにすることもできるし、ファイル名を指定することによりファイル単位に行うようにすることもできる。後者の場合にはストレージ装置 200 は、情報処理装置 100 からのファイルレベルでのアクセスを実現する N A S (Network Attached Storage) として機能する。

【0029】

ディスク I / F 214 は記憶ボリュームに対するデータ入出力制御を行うための通信インタフェースとして機能する。

内部 I / F 215 は管理用装置 215 と通信を行うための通信インタフェースとして機能する。またストレージ装置 200 には複数の制御装置 210 が備えられるようにすることができる。この場合には、内部 I / F 215 は他の制御装置 210 との間で通信を行うための通信インタフェースとして機能するようにすることもできる。

【0030】

=== 管理用装置 ===

図 1 に戻って、管理用装置 250 はストレージ装置 200 を保守・管理するために用いられるコンピュータである。管理用装置 250 を操作することにより、

例えばディスクドライブの設定や、記憶ボリュームの設定、制御装置 210 において実行されるスナップショット制御プログラム 800 のインストール等を行うことができる。また管理用装置 250 は記録媒体読み取り装置を備えており、記録媒体読み取り装置を用いて、スナップショット制御プログラム 800 が記録された記録媒体からスナップショット制御プログラム 800 を読み取るようにすることもできる。

【0031】

またスナップショットコマンドは管理用装置 250 からストレージ装置 200 に送信されるようにすることもできる。管理用装置 250 はストレージ装置 200 に外付けされている形態とすることもできるし、内蔵されている形態とすることもできる。また管理用装置 250 が備える機能を情報処理装置 100 が備えるようにすることもできる。

【0032】

===スナップショット管理テーブル、S-VOL変換テーブル===

次にスナップショット管理テーブル 400 について図 5 を用いて説明する。スナップショット管理テーブル 400 は、上述したスナップショット制御を行うために用いられるテーブルである。情報処理装置 100 からセカンダリボリューム 230 に対してデータ入出力要求があった場合に、CPU 211 がスナップショット管理テーブル 400 を参照することにより、そのデータ入出力要求に対して、プライマリボリューム 220 にアクセスすべきなのか、差分ボリューム 240 にアクセスすべきなのかを知ることができる。スナップショット管理テーブル 400 は、図 5 に一例として示すビットマップに限らず、リスト構造や、物理アドレスを示すポインタ等とすることもできる。スナップショット管理テーブル 400 は、管理情報または制御情報として、以下に説明するようなプライマリボリューム 220 とセカンダリボリューム 230 との対応付けの管理又は制御のために用いることができるものであれば、どのような態様とすることもできる。

【0033】

スナップショット管理テーブル 400 において、プライマリボリューム BM (P) 410 はプライマリボリューム 220 のデータの保持状態を示し、セカンダリ

ボリュームBM (S) 420はセカンダリボリューム230のデータの保持状態を示す。具体的には、プライマリボリュームBM (P) 410又はセカンダリボリュームBM (S) 420のそれぞれにおいて、ビットの値が” 0” の場合は、当該データが有効であることを示す。またビットの値が” 1” の場合は、当該データが無効であることを示す。

【0034】

なお、ビットマップ上の各ビットの指定は、プライマリボリュームBM (P) 410においては $(x p i, y p j)$ ($i, j = 1, 2, \dots$)、セカンダリボリュームBM (S) 420においては $(x s i, y s j)$ ($i, j = 1, 2, \dots$) を指定することにより行われる。

【0035】

プライマリボリュームBM (P) 410及びセカンダリボリュームBM (S) 420のそれぞれのビットマップ上の対応するビットは、プライマリボリューム220及びセカンダリボリューム230のそれぞれ対応する記憶領域に対応する。プライマリボリュームBM (P) 410、及びセカンダリボリュームBM (s) 420のそれぞれのビットの値は、プライマリボリューム220及びセカンダリボリューム230上のデータの記憶単位であるブロック、ないしトラックなどのデータの状態に対応する。例えば各ビットが1トラックに対応する場合、 $(x p i, y p j)$ で特定されるビットの値は、例えばプライマリボリューム220のシリンダ番号 $x p i$ 、ヘッド番号 $y p j$ で一意に定められるトラックのデータの状態を示す。なお、セカンダリボリューム230は仮想的な記憶ボリュームであるので、実際のデータはプライマリボリューム220や差分ボリューム240、他の記憶ボリューム241に記憶されている。そのためセカンダリボリューム230に対するデータ入出力要求があった場合には、図6に示すS-VOL変換テーブルにより、セカンダリボリューム230における仮想的な記憶領域を示す $(x s i, y s j)$ ($i, j = 1, 2, \dots$) をプライマリボリューム220、差分ボリューム240、又は他の記憶ボリューム241における実際の記憶領域を示す $(x s i', y s j')$ ($i', j' = 1, 2, \dots$) に変換して、データ入出力処理が行われる。

【0036】

上述したように、プライマリボリュームBM (P) 410、又はセカンダリボリュームBM (S) 420において、ビットの値が” 0” の場合は当該データが有効であることを示し、ビットの値が” 1” の場合は当該データが無効であることを示す。例えば図5に示すように、プライマリボリュームBM (P) 410上のビットの値が全て” 0” であり、セカンダリボリュームBM (S) 420上のビットの値が全て” 1” である場合には、情報処理装置100からセカンダリボリューム230にデータの読み出し要求があると、ストレージ装置200は、プライマリボリューム220からデータを読み出して情報処理装置100に送信する。つまりストレージ装置200は、プライマリボリュームBM (P) 410上のビットの値を全て” 0” とし、セカンダリボリュームBM (S) 420上のビットの値を全て” 1” とすることにより、プライマリボリューム220の複製をセカンダリボリューム230に記憶しているかのように制御を行うことができる。

【0037】

一方、ストレージ装置200が情報処理装置100からスナップショットコマンドを受信し、ある時点のプライマリボリューム220のデータの複製がセカンダリボリューム230に記憶されているかのように制御する場合には、ある時点の後に情報処理装置100からプライマリボリューム220のデータに対する更新要求を受信すると、更新前のある時点のデータを差分ボリューム240に記憶し、更新後のデータをプライマリボリューム220に記憶すると共に、プライマリボリュームBM (P) 410、及びセカンダリボリュームBM (S) 420のビットの値をどちらも” 0” に更新する。この場合にプライマリボリューム220に対するデータ読み出し要求があった場合には、プライマリボリューム220から更新後のデータが読み出され、セカンダリボリューム230に対するデータ読み出し要求があった場合には、プライマリボリューム220からではなく、差分ボリューム240または他の記憶ボリューム241から更新前のある時点のデータが読み出される。

【0038】

もちろん、上記のようにある時点のプライマリボリューム220のデータの複

製がセカンダリボリューム 230 に記憶されているかのように制御する場合であっても、プライマリボリューム 220 のデータに対する更新要求を受信していない場合には、ある時点のデータはプライマリボリューム 220 に記憶されたままであるので、プライマリボリューム BM (P) 410 の当該データに対するビットの値は” 0 ”であり、セカンダリボリューム BM (S) 420 の当該データに対するビットの値は” 1 ”である。従って、この場合に情報処理装置 100 からセカンダリボリューム 230 にデータの読み出し要求があった場合には、ストレージ装置 200 は、プライマリボリューム 220 からデータを読み出して情報処理装置 100 に送信する。

【0039】

このように、ストレージ装置 200 は、スナップショット管理テーブル 400 を参照することにより、情報処理装置 100 からセカンダリボリューム 230 にデータ入出力要求があった場合に、プライマリボリューム 220 からデータを読み出すべきなのか、差分ボリューム 240 からデータを読み出すべきなのかを知ることができる。

【0040】

なおスナップショット管理テーブル 400 は、制御装置 210 内のメモリ 212 上に設けられる制御情報として重要であり、メモリ 212 が揮発性メモリの場合は、不揮発性メモリにスナップショット管理テーブル 400 のバックアップを取っておくようにすることもできる。またスナップショット管理テーブル 400 をストレージ装置 200 が備える記憶ボリュームに記憶するようにすることもできる。後述するその他の各テーブルについても同様である。

【0041】

=== ボリューム管理テーブル、差分ボリューム管理テーブル ===

次に、本実施の形態に係るボリューム管理テーブル 600 について図 9 を用いて説明する。ボリューム管理テーブル 600 は、ストレージ装置 200 が備える記憶ボリュームを管理するためのテーブルである。

【0042】

ボリューム管理テーブル 600 は、” LUN ” 欄、” 記憶容量 ” 欄、” 空き容

量”欄、“スナップショット用ボリューム管理”欄、“用途”欄を備える。”スナップショット用ボリューム管理”欄は、さらに”ボリューム種別”欄、“ペアボリュームNO”欄、“差分ボリュームNO”欄、“差分ボリューム使用可否”欄を備える。

”LUN”欄は、記憶ボリュームを識別するための番号である。例えば管理用装置250等を用いて記憶ボリュームを設定する際に付与される。

”記憶容量”欄は、記憶ボリュームの総記憶容量を示す。”記憶容量”欄に「-」と記載されている記憶ボリュームは、ストレージ装置200が仮想的に備える記憶ボリュームであることを示す。

”空き容量”欄は、記憶ボリュームの未使用記憶領域の記憶容量を示す。”記憶容量”欄に「-」と記載されている記憶ボリュームは、“空き容量”欄も「-」と記載されている。

【0043】

”ボリューム種別”欄は、当該記憶ボリュームが、スナップショット制御においてプライマリボリューム220として用いられるのか、セカンダリボリューム230として用いられるのか、差分ボリューム240として用いられるのかを示す。「P」と記載されている場合はプライマリボリューム220として用いられる記憶ボリュームであることを示す。「S」と記載されている場合はセカンダリボリューム230として用いられる記憶ボリュームであることを示す。「差分」と記載されている場合は差分ボリューム240として用いられる記憶ボリュームであることを示す。「-」と記載されている場合はスナップショット制御の対象ではない記憶ボリュームであることを示す。なお上述したように、セカンダリボリューム230はストレージ装置200が仮想的に備える記憶ボリュームであるから、“ボリューム種別”欄が「S」である記憶ボリュームについては、“記憶容量”欄、“空き容量”欄には「-」と記されている。

【0044】

”ペアボリュームNO”欄には、プライマリボリューム220またはセカンダリボリューム230の、それぞれのセカンダリボリューム230またはプライマリボリューム220のLUNが記載される。

【0045】

”差分ボリュームNO”欄にはプライマリボリューム220に対する差分ボリューム240のLUNが記載される。図9に示す例では、LUN=1のプライマリボリューム220に対する差分ボリューム240はLUN=6である。またLUN=3のプライマリボリューム220に対する差分ボリューム240はLUN=5及びLUN=9である。ここでLUN=5の”空き容量”欄に着目すると、0GB (Giga Byte) と記載されている。つまり差分ボリューム240の記憶領域を使いきってしまっている状態であることが分かる。そのため、LUN=9の未使用記憶領域を差分ボリューム240として使用することにより、プライマリボリューム220 (LUN=3) のある時点のデータの記憶を継続して行うようにしている。

【0046】

”差分ボリューム使用可否”欄は、当該記憶ボリュームについて差分ボリューム240としての使用を許可するかどうかを示す欄である。「○」と記載されている場合は、当該記憶ボリュームは差分ボリューム240として使用されることが許可されていることを表す。「－」と記載されている場合には、当該記憶ボリュームは差分ボリューム240として使用されることが許可されていないことを表す。例えばLUN=6の空き容量は現在は100GBあるが、将来0GBになった場合には他の記憶ボリューム241の未使用記憶領域を差分ボリューム240として使用することになる。その際”差分ボリューム使用可否”欄に「○」が記載されている記憶ボリュームの中から他の記憶ボリューム241が選定される。

【0047】

なお差分ボリューム240として使用する他の記憶ボリューム241の選定は、図10に示す差分ボリューム管理テーブル700を参照して行うようにすることもできる。差分ボリューム管理テーブル700は、プライマリボリューム220毎に差分ボリューム240の候補となる記憶ボリュームが記載されたものである。差分ボリューム管理テーブル700は、例えば管理用装置250が備えるユーザインタフェースからオペレータ等により入力された差分ボリューム240の

候補を指定するための情報に応じて作成されるようにすることができる。

【0048】

差分ボリューム 2 4 0 として使用する他の記憶ボリューム 2 4 1 の選定は、差分ボリューム管理テーブル 7 0 0 を参照して行う場合であっても参照せずに行う場合であっても、例えば候補となる記憶ボリュームの中から、空き容量に応じて記憶ボリュームを選定するようにすることもできるし、L U N に応じて記憶ボリュームを選定するようにすることもできる。また単位時間当たりのデータ入出力要求の回数に応じて記憶ボリュームを選定するようにすることもできるし、空き容量の総記憶容量に対する割合に応じて記憶ボリュームを選定するようにすることもできる。さらには、例えば管理用装置 2 5 0 が備えるユーザインタフェースから、オペレータ等が差分ボリューム 2 4 0 として使用する記憶ボリュームを指定するための情報を入力することにより、記憶ボリュームを選定するようにすることもできる。

【0049】

”用途”欄は、各記憶ボリュームの使用用途を示す。「DB」と記載されている場合はデータベースアプリケーションにより使用される記憶ボリュームであることを示す。「WEB」と記載されている場合はWEBアプリケーションにより使用される記憶ボリュームであることを示す。「FS」と記載されている場合はオペレーティングシステムにより使用される記憶ボリュームであることを示す。

「-」と記載されている場合は特に特定のアプリケーションに使用されている記憶ボリュームではないことを示す。例えば L U N = 5 は差分ボリューム 2 4 0 としてのみ使用されている。なお、”用途”欄に例えば「DB」と記載されている場合であっても、データベースアプリケーション以外の他のアプリケーションと共用されているようにすることもできる。また各記憶ボリュームは上記以外の用途に使用されるようにすることもできる。

【0050】

図 9 に示すボリューム管理テーブル 6 0 0 に記載される内容や、図 1 0 に示す差分ボリューム管理テーブル 7 0 0 に記載される内容は、管理用装置 2 5 0 や情報処理装置 1 0 0 が備えるユーザインタフェースに表示するようにすることができ

る。例えば差分ボリューム 240 の空き容量が表示されることにより、オペレータ等は差分ボリューム 240 の未使用記憶領域の記憶容量を把握することができる。これにより、例えば差分ボリューム 240 が不足してきたと判断した場合には、差分ボリューム 240 の記憶領域を使いきってしまう前にディスクドライブを増設して、差分ボリューム 240 の記憶容量を増やすようにすることもできる。またストレージ装置 200 が備える記憶ボリュームの設定を変更して差分ボリューム 240 の記憶容量を増やすようにすることもできる。

【0051】

====プライマリボリュームからのデータの読み出し処理====

次に、本実施の形態に係るプライマリボリューム 220 からデータを読み出す場合の処理のフローチャートを図 11 に示す。なお以下の処理は、本実施の形態に係る各種の動作を行うためのコードから構成されるスナップショット制御プログラム 800 を、CPU 211 が実行することにより実現される。以下に説明するその他の処理についても同様である。

【0052】

まずストレージ装置 200 は情報処理装置 100 からプライマリボリューム 220 に対するデータの読み出し要求を受信する。そうすると、まずスナップショット管理テーブル 400 をチェックし、目的のデータをプライマリボリューム 220 から読み出すべきなのか、セカンダリボリューム 230 から読み出すべきなのかを判定する (S1000)。プライマリボリューム BM (P) 410 上のビットの値が” 0 ”であり、セカンダリボリューム BM (S) 420 上のビットの値が” 1 ”である場合には、プライマリボリューム 220 に記憶されているデータが有効なデータであるので、S1001において” YES ”に進む。そしてプライマリボリューム 220 から目的のデータを読み出して情報処理装置 100 に送信する (S1002)。

【0053】

一方、プライマリボリューム BM (P) 410 上のビットの値が” 1 ”であり、セカンダリボリューム BM (S) 420 上のビットの値が” 0 ”である場合には、セカンダリボリューム 230 に記憶されているデータが有効なデータであるので

、S1001において” NO” に進む。そしてS-VOL変換テーブル500を参照して目的のデータの記憶位置を取得し（S1003）、差分ボリューム240からデータを読み出して情報処理装置100に送信する（S1004）。

【0054】

===プライマリボリュームへのデータの書き込み処理===

次に、本実施の形態に係るプライマリボリューム220へデータを書き込む場合の処理のフローチャートを図12に示す。

まずストレージ装置200は情報処理装置100からプライマリボリューム220に対するデータの書き込み要求を受信する。そしてスナップショット管理テーブル400をチェックし（S2000）、プライマリボリューム220に記憶されている更新前のデータをセカンダリボリューム230に反映する必要があるのか、ないのかを判定する。例えば、ストレージ装置200がプライマリボリューム220に記憶されているデータと同一内容のデータがセカンダリボリューム230から読み出されるように制御している場合には、更新前のプライマリボリューム220に記憶されているデータをセカンダリボリューム230に反映する必要はない。また、ストレージ装置200がある時点のプライマリボリューム220に記憶されているデータがセカンダリボリューム230から読み出されるように制御している場合には、ある時点のデータがセカンダリボリューム230に未反映であれば反映する必要があるし、反映済みであれば反映する必要はない。

【0055】

プライマリボリューム220のある時点のデータをセカンダリボリューム230に反映する必要がない場合には、S2001において” YES” に進む。そしてプライマリボリューム220にデータを書き込む（S2002）。

一方、プライマリボリューム220のある時点のデータをセカンダリボリューム230に反映する必要がある場合には、S2001において” NO” に進む。そしてボリューム管理テーブル600を参照し、差分ボリューム240の未使用記憶領域の記憶容量すなわち空き容量をチェックする（S2003）。

【0056】

空き容量が規定容量（判定値）以上ある場合には、” YES” に進む（S2004）。

そして当該データのセカンダリボリューム 2 3 0 における記憶位置と、差分ボリューム 2 4 0 における記憶位置とを S-VOL 変換テーブル 5 0 0 に登録し (S2005)、差分ボリューム 2 4 0 にある時点のプライマリボリューム 2 2 0 のデータの複製を書き込む (S2006)。なお、S2004における空き容量と規定容量との比較において、例えば空き容量が規定容量よりも大きい場合に” YES” に進むようにすることもできる。次に、スナップショット管理テーブル 4 0 0 のプライマリボリューム BM (P) 4 1 0 及びセカンダリボリューム BM (S) 4 2 0 のビットの値を共に” 0” に更新し (S2007)、プライマリボリューム 2 2 0 にデータを書き込む (S2002)。

【 0 0 5 7 】

S2004において、差分ボリューム 2 4 0 の空き容量が規定容量以上ない場合には、” NO” に進む。この場合は、ある時点のプライマリボリューム 2 2 0 のデータを記憶する差分ボリューム 2 4 0 として使用するための他の記憶ボリューム 2 4 1 を選定することが必要になる。選定は、上述したように例えばボリューム管理テーブル 6 0 0 や差分ボリューム管理テーブル 7 0 0 を参照することにより行うことができる (S2008)。差分ボリューム 2 4 0 として使用するための他の記憶ボリューム 2 4 1 が選定できた場合には、S2009において” YES” に進む。そしてボリューム管理テーブル 6 0 0 の各記憶ボリュームについての情報を更新して (S2010)、S2003に戻る。その後はS2004にて” YES” に進み、上述した各ステップの処理を行って (S2005乃至S2007)、プライマリボリューム 2 2 0 にデータを書き込む (S2002)。

【 0 0 5 8 】

なおS2006において、上記選定された他の記憶ボリューム 2 4 1 の未使用記憶領域にある時点のプライマリボリューム 2 2 0 のデータの複製を書き込む場合には、当該未使用記憶領域のうちデータアクセス性能の低い記憶領域に書き込むようにすることもできる。例えばディスクドライブなどでは、ディスクの外周側のトラックほど、1トラックあたりのセクタの数が多いものがある。この場合ディスクの外周側と内周側とではディスクが1回転する時間の間に読み書きすることができるデータ量が異なり、ディスクの内周側ほどデータアクセス性能が低いこ

となる。ディスクの外周側のようなデータアクセス性能の高い記憶領域は、例えば情報処理装置 100 が情報処理サービスを提供するためのデータを記憶するために使用するようし、ディスクの内周側のようなデータアクセス性能の低い記憶領域は、本実施の形態におけるプライマリボリューム 220 のある時点のデータを記憶するために使用するようにする。これにより、他の記憶ボリューム 241 の未使用記憶領域にある時点のプライマリボリューム 220 のデータの複製を書き込むようにしても、他の記憶ボリューム 241 を用いて提供される本来の情報処理サービスに与える影響を抑制することができる。なお、記憶ボリュームのうちデータアクセス性能の低い記憶領域は、例えば上述したトラック番号により識別することができる。またデータの記憶アドレスやセクタ番号等によっても識別することができる。

【0059】

一方、S2008において差分ボリューム 240 として使用するための他の記憶ボリューム 241 が選定できなかった場合には、S2009において” NO” に進む。差分ボリューム 240 として使用するための他の記憶ボリューム 241 が選定できなかった場合とは、例えば差分ボリューム 240 として使用可能な記憶ボリュームがボリューム管理テーブル 600 に登録されていない場合や、差分ボリューム管理テーブル 700 に差分ボリューム候補が登録されていない場合、あるいは差分ボリューム 240 として使用可能な全ての記憶ボリュームを既に差分ボリューム 240 として使用してしまっている場合などである。続いて S2011 にて、差分ボリューム 240 の未使用記憶領域に、ある時点のプライマリボリューム 220 のデータをまだ書き込めるだけの記憶容量が残っているかどうかを判定し、まだ残っている場合には” NO” に進む。そして例えば管理用装置 250 が備えるユーザインタフェースに警告を表示する (S2012)。警告が表示されることにより、ストレージ装置 200 を管理するオペレータは、差分ボリューム 240 を使い切ってしまう前に対策を行うことが可能となる。例えばディスクドライブを増設して差分ボリューム 240 の記憶容量を増やすようにすることもできる。あるいはスナップショットコマンドをストレージ装置 200 に送信して、ある時点のプライマリボリューム 220 のデータの維持を中止するようにすることもできる。

【0060】

S2011にて、差分ボリューム240の未使用記憶領域に、ある時点のプライマリボリューム220のデータをまだ書き込めるだけの記憶容量が残っていない場合には”YES”に進む。そして例えば管理用装置250が備えるユーザインタフェースに警告を表示し(S2013)、ある時点のプライマリボリューム220のデータの維持を中止する(S2014)。これにより差分ボリューム240に記憶されているある時点のプライマリボリューム220のデータが無効とされる。スナップショット管理テーブル400のプライマリボリュームBM(P)410のビットが”0”に更新され、セカンダリボリュームBM(S)420のビットの値が”1”に更新される。そしてプライマリボリューム220にデータを書き込む(S2002)。

【0061】

===セカンダリボリュームからのデータの読み出し処理===

次に、本実施の形態に係るセカンダリボリューム230からデータを読み出す場合の処理のフローチャートを図13に示す。

まずストレージ装置200は情報処理装置100からセカンダリボリューム230に対するデータの読み出し要求を受信する。そうすると、まずスナップショット管理テーブル400をチェックし、目的のデータをプライマリボリューム220から読み出すべきなのか、セカンダリボリューム230から読み出すべきなのかを判定する(S3000)。プライマリボリュームBM(P)410上のビットの値が”0”であり、セカンダリボリュームBM(S)420上のビットの値が”1”である場合には、プライマリボリューム220に記憶されているデータが有効なデータであるので、S3001において”NO”に進む。そしてスナップショット管理テーブル400を参照して、セカンダリボリューム230の当該データに対応するプライマリボリューム220の記憶位置を特定し(S3004)、プライマリボリューム220から目的のデータを読み出して情報処理装置100に送信する(S3005)。

【0062】

一方、プライマリボリュームBM(P)410上のビットの値が”1”であり、

セカンダリボリュームBM (S) 4 2 0 上のビットの値が” 0 ” である場合、又はどちらの値も” 0 ” であるには、セカンダリボリューム 2 3 0 に記憶されているデータが有効なデータであるので、S3001において” YES ” に進む。そして S - V O L 変換テーブル 5 0 0 を参照して目的のデータの記憶位置を取得し (S3002)、差分ボリューム 2 4 0 からデータを読み出して情報処理装置 1 0 0 に送信する (S3003)。

【 0 0 6 3 】

===セカンダリボリュームへのデータの書き込み処理===

次に、本実施の形態に係るセカンダリボリューム 2 3 0 へデータを書き込む場合の処理のフローチャートを図 1 4 に示す。

まずストレージ装置 2 0 0 は情報処理装置 1 0 0 からセカンダリボリューム 2 3 0 に対するデータの書き込み要求を受信する。そしてスナップショット管理テーブル 4 0 0 をチェックし (S4000)、プライマリボリューム 2 2 0 に記憶されている更新前のデータがセカンダリボリューム 2 3 0 に反映済みであるのか、未反映であるのかを判定する。例えば、ある時点のデータがセカンダリボリューム 2 3 0 に未反映であれば反映する必要があるし、反映済みであれば反映する必要はない。

【 0 0 6 4 】

プライマリボリューム 2 2 0 のある時点のデータをセカンダリボリューム 2 3 0 に反映済みの場合には、S4001において” YES ” に進む。そしてセカンダリボリューム 2 3 0 にデータを書き込む (S4002)。実際には、S - V O L 変換テーブル 5 0 0 に記憶されているセカンダリボリューム 2 3 0 に対応する差分ボリューム 2 4 0 における記憶位置にデータを書き込む (S4002)。

一方、プライマリボリューム 2 2 0 のある時点のデータがセカンダリボリューム 2 3 0 に未反映である場合には、S4001において” NO ” に進む。そしてボリューム管理テーブル 6 0 0 を参照し、差分ボリューム 2 4 0 の未使用記憶領域の記憶容量すなわち空き容量をチェックする (S4003)。

【 0 0 6 5 】

空き容量が規定容量 (判定値) 以上ある場合には、” YES ” に進む (S4004)。

そして当該データのセカンダリボリューム 2 3 0 における記憶位置と、差分ボリューム 2 4 0 における記憶位置とを S-VOL 変換テーブル 5 0 0 に登録し (S4005)、差分ボリューム 2 4 0 にある時点のプライマリボリューム 2 2 0 のデータの複製を書き込む (S4006)。なお、S4004 における空き容量と規定容量との比較において、例えば空き容量が規定容量よりも大きい場合に” YES” に進むようにすることもできる。次にスナップショット管理テーブル 4 0 0 のプライマリボリューム BM (P) 4 1 0 及びセカンダリボリューム BM (S) 4 2 0 のビットの値を共に” 0” に更新し (S4007)、セカンダリボリューム 2 3 0、すなわち差分ボリューム 2 4 0 にデータを書き込む (S4002)。

【 0 0 6 6 】

S4004 において、差分ボリューム 2 4 0 の空き容量が規定容量以上ない場合には、” NO” に進む。この場合は、ある時点のプライマリボリューム 2 2 0 のデータを記憶する差分ボリューム 2 4 0 として使用するための他の記憶ボリューム 2 4 1 を選定することが必要になる。選定は、上述したように例えばボリューム管理テーブル 6 0 0 や差分ボリューム管理テーブル 7 0 0 を参照することにより行うことができる (S4008)。差分ボリューム 2 4 0 として使用するための他の記憶ボリューム 2 4 1 が選定できた場合には、S4009 において” YES” に進む。そしてボリューム管理テーブル 6 0 0 の各記憶ボリュームについての情報を更新して (S4010)、S4003 に戻る。その後は S4004 にて” YES” に進み、上述した各ステップの処理を行って (S4005 乃至 S4007)、セカンダリボリューム 2 3 0 すなわち差分ボリューム 2 4 0 にデータを書き込む (S4002)。

【 0 0 6 7 】

なお S4006 において、上記選定された他の記憶ボリューム 2 4 1 の未使用記憶領域にある時点のプライマリボリューム 2 2 0 のデータの複製を書き込む場合には、当該未使用記憶領域のうちデータアクセス性能の低い記憶領域に書き込むようにすることもできる。

【 0 0 6 8 】

一方、S4008 において差分ボリューム 2 4 0 として使用するための他の記憶ボリューム 2 4 1 が選定できなかった場合には、S4009 において” NO” に進む。続

いてS4011にて、差分ボリューム240の未使用記憶領域に、ある時点のプライマリボリューム220のデータをまだ書き込めるだけの記憶容量が残っているかどうかを判定し、まだ残っている場合には”NO”に進む。そして例えば管理用装置250が備えるユーザインタフェースに警告を表示する(S4012)。警告が表示されることにより、ストレージ装置200を管理するオペレータは、差分ボリューム240を使い切ってしまう前に対策を行うことが可能となる。例えばディスクドライブを増設して差分ボリューム240の記憶容量を増やすようにすることもできる。あるいはスナップショットコマンドをストレージ装置200に送信して、ある時点のプライマリボリューム220のデータの維持を中止するようにすることもできる。

【0069】

S4011において、差分ボリューム240の未使用記憶領域に、ある時点のプライマリボリューム220のデータをまだ書き込めるだけの記憶容量が残っていない場合には”YES”に進む。そして例えば管理用装置250が備えるユーザインタフェースに警告を表示し(S4013)、ある時点のプライマリボリューム220のデータの維持を中止する(S4014)。

【0070】

===記憶ボリュームへのデータの書き込みについて===

一般に、ストレージ装置200が備える記憶ボリュームへデータを書き込むためには、当該記憶ボリュームに対するフォーマット処理を行っておく必要がある。フォーマット処理とは、記憶ボリュームを構成するディスクドライブ等の記憶メディアにより提供される記憶領域を、データを記憶するための区画に区分することにより、データの書き込みや読み出しを可能とするための処理である。区画とは、例えばディスクドライブに設定されるトラック上に設定されるセクタとすることができる。

【0071】

本実施の形態に係るストレージ装置200においては、フォーマット処理として2種類の方式を実行可能である。すなわち図17に示すように、第1のフォーマット方式と第2のフォーマット方式とである。なお第1のフォーマット方式で

フォーマット処理を行うか、第2のフォーマット方式でフォーマット処理を行うかは、例えば管理用装置250が備えるユーザインタフェースからの指示により、切り替えるようにすることができる。もちろん、本実施の形態に係るストレージ装置200は、例えば第2のフォーマット方式のみを実行可能とすることもできる。

【0072】

第1のフォーマット方式では、ストレージ装置200は、情報処理装置100や管理用装置250からフォーマット処理実行の指示を受信すると、指定された記憶ボリュームに対するフォーマット処理を開始する。しかしフォーマット処理を行っている間はその記憶ボリュームに対するデータ入出力要求を受け付けない。フォーマット処理が完了した後にその記憶ボリュームに対するデータ入出力要求の受け付けを開始する。

【0073】

一方第2のフォーマット方式では、ストレージ装置200は、情報処理装置100や管理用装置250からフォーマット処理実行の指示を受信すると、指定された記憶ボリュームに対するフォーマット処理を開始するが、フォーマット処理を行っている間であってもその記憶ボリュームに対するデータ入出力要求を受け付ける。この場合、ストレージ装置200はフォーマット処理が完了していなくても、当該記憶ボリュームに対するデータ入出力要求を受信した場合には、あたかもフォーマット処理が完了しているかのようにデータ入出力制御を行う。

【0074】

このため、第2のフォーマット方式でフォーマット処理を行っている間にデータ入出力要求を受け付けた場合には、まだフォーマット処理が済んでいない記憶領域に対してデータ入出力処理を行わなければならない場合がある。例えば本実施の形態において、差分ボリューム240のフォーマット処理が第2のフォーマット方式で実行される場合には、図12及び図14において説明した差分ボリューム240へのデータの書き込みの際に、当該データが書き込まれる予定の記憶領域のフォーマット処理が完了していない場合がある。この場合は図18に示すように、当該データを書き込むために必要な記憶領域についてフォーマット処理

を先に行ってからデータを書き込むようにする。これにより、フォーマット処理がまだ完了していない記憶ボリュームに対してデータ入出力要求があった場合であってもデータ入出力処理を行うことが可能となる。

【0075】

以下、本実施の形態に係る第2のフォーマット方式について説明する。

第2のフォーマット方式により記憶ボリュームのフォーマット処理を行う場合には、上述した区画毎にフォーマット処理が済んでいるか、済んでいないかをフォーマット管理テーブル900を用いて管理する。フォーマット管理テーブル900を用いてフォーマット処理が行われる様子を図15に示す。フォーマット管理テーブル900は記憶ボリューム毎に設けられるようにすることができる。

【0076】

フォーマット管理テーブル900の一例を図16に示す。図16において $(x f i, y f j)$ ($i, j = 1, 2, \dots$) は記憶ボリューム上の記憶位置を区画毎に特定するための情報である。例えば各区画がディスクドライブ上の1セクタに対応する場合には、 $x f i$ ($i = 1, 2, \dots$) は例えばディスクドライブ上に設定されるトラック番号に対応し、 $y f j$ ($j = 1, 2, \dots$) はトラック上に設定されるセクタ番号に対応するようにすることができる。そしてフォーマット処理が済んでいる区画に対しては、フォーマット管理テーブル900において“1”が記載される。フォーマット処理が済んでいない区画に対しては“0”が記載される。これによりストレージ装置200は、記憶ボリュームの各区画についてフォーマット処理が済んでいるのか、済んでいないのかを管理することができる。そのためフォーマット処理中であっても情報処理装置100からデータ入出力要求を受け付けることができ、フォーマット処理が済んでいない記憶領域に対するデータ入出力要求であってもデータ入出力処理を行うようにすることが可能となる。

【0077】

次に第2のフォーマット形式でフォーマット処理が行われている記憶ボリュームに対してデータの書き込み要求があった場合の処理の流れについて、図19に示すフローチャートを用いて説明する。なお図19においては、上述した差分ボ

リユーム 2 4 0 に対して第 2 のフォーマット方式でフォーマット処理が行われている場合について記載されているが、他の記憶ボリユームに対して第 2 のフォーマット方式でフォーマット処理が行われている場合も同様である。また以下の処理は、本実施の形態に係る各種の動作を行うためのコードから構成されるスナップショット制御プログラム 8 0 0 を、CPU 2 1 1 が実行することにより実現される。もちろん、CPU 2 1 1 はスナップショット制御プログラム 8 0 0 の他に、他のプログラムを実行することにより以下の処理を実現するようにすることもできる。

【0 0 7 8】

まずストレージ装置 2 0 0 は、情報処理装置 1 0 0 からデータの書き込み要求を受信する (S5000)。次にデータ書き込み要求の対象となっている記憶ボリユームについてフォーマット処理が完了しているかどうかをチェックする (S5001)。このチェックは、例えばフォーマット処理が完了しているかどうかを示すフラグを記憶ボリユーム毎に設けて、このフラグをチェックすることにより行うようにすることができる。この場合、フラグは記憶ボリユームに対するフォーマット処理を完了したときに、ストレージ装置 2 0 0 によりセットされる。次に、当該記憶ボリユームに対するフォーマット処理が完了している場合には”完了”に進み、差分ボリユーム 2 4 0 にデータを書き込む (S5002)。

【0 0 7 9】

フォーマット処理が完了していない場合には”未完”に進む。そしてフォーマット管理テーブル 9 0 0 を参照して、当該データ入出力要求の対象となっている記憶領域についてのフォーマットが済んでいるかどうかをチェックする (S5003)。フォーマット処理が完了している場合には”1”に進み、差分ボリユーム 2 4 0 にデータを書き込む (S5004)。

【0 0 8 0】

フォーマット処理が完了していない場合には、”0”に進む。そしてまず当該記憶領域に対するフォーマット処理を行うために、データ入出力要求の処理を一時的に停止する (S5005)。その間に当該記憶領域に対するフォーマット処理を行う (S5006)。そしてフォーマット管理テーブル 9 0 0 の更新を行った後 (S50

07)、データ入出力要求の処理を再開し、差分ボリューム 2 4 0 の当該記憶領域にデータを書き込む (S5008)。

【0 0 8 1】

これにより、第 2 のフォーマット方式によってフォーマット処理を行っている間であってもデータ入出力要求を受信することができ、フォーマット処理が済んでいない記憶領域に対するデータ入出力要求であってもデータ入出力処理を行うようにすることができる。

【0 0 8 2】

以上説明したように、本実施の形態によれば、スナップショット制御に用いられる差分ボリューム 2 4 0 の記憶容量が不足した場合であっても、スナップショット制御を継続することが可能となる。

【0 0 8 3】

これにより、例えばあらかじめ差分ボリューム 2 4 0 として用意しておく記憶容量を抑制することが可能となる。あらかじめ用意しておく差分ボリューム 2 4 0 を無くすようにすることも可能である。差分ボリューム 2 4 0 はプライマリボリューム 2 2 0 の更新前のある時点のデータを記憶するための記憶ボリュームであり、情報処理装置 1 0 0 により提供される情報処理サービスのために使用される記憶ボリュームではない。このため、本実施の形態により、差分ボリューム 2 4 0 として用意しておく記憶容量を抑制可能となったことにより、コンピュータシステム及びストレージ装置 2 0 0 のコストパフォーマンスを向上させることが可能となる。

【0 0 8 4】

また、例えばバックアップ処理の実行中に差分ボリューム 2 4 0 の記憶領域を使い切ってしまった場合であっても、バックアップ処理を中止する必要がなくなる。これによりバックアップ処理を確実に行えるようにできるため、コンピュータシステムの信頼性を向上させることができる。

【0 0 8 5】

同様に、例えばセカンダリボリューム 2 3 0 に仮想的に記憶されている、ある時点のプライマリボリューム 2 2 0 のデータを用いて、テストプログラムの実行

やシミュレーションプログラムの実行を行う場合にも、プログラムの実行中に差分ボリューム 2 4 0 の記憶領域を使い切ってしまったとしても、プログラムの実行を継続することが可能となる。

【 0 0 8 6 】

また本実施の形態においては、差分ボリューム 2 4 0 の空き容量が管理用装置 2 5 0 や情報処理装置 1 0 0 が備えるユーザインタフェースに表示される。これによりコンピュータシステムを管理するオペレータは、スナップショット制御の実施状況を把握することができ、例えば差分ボリューム 2 4 0 の増設等の処置を的確に行うことが可能となる。

【 0 0 8 7 】

また差分ボリューム 2 4 0 の空き容量が少なくなった場合には、管理用装置 2 5 0 や情報処理装置 1 0 0 が備えるユーザインタフェースに警告が表示される。これによりコンピュータシステムを管理するオペレータは、差分ボリューム 2 4 0 を使い切って突然スナップショット制御が中止されてしまう前に、例えばディスクドライブを増設して差分ボリューム 2 4 0 の記憶容量を増やす等の対策や、スナップショット制御の中止等の処置を行うことが可能となる。

【 0 0 8 8 】

また本実施の形態においては、記憶ボリュームに対するフォーマット処理を行っている間であっても、その記憶ボリュームに対するデータ入出力要求を受け付けることができる。これにより、フォーマット処理中であっても情報処理装置 1 0 0 からデータ入出力要求を受け付け、フォーマット処理が済んでいない記憶領域に対するデータ入出力要求であってもデータ入出力処理を行うようにすることが可能となる。

【 0 0 8 9 】

以上本実施の形態について説明したが、上記実施例は本発明の理解を容易にするためのものであり、本発明を限定して解釈するためのものではない。本発明はその趣旨を逸脱することなく変更、改良され得ると共に、本発明にはその等価物も含まれる。

【 0 0 9 0 】

【発明の効果】

ストレージ装置、及びストレージ装置の制御方法を提供することができる。

【図面の簡単な説明】

【図 1】 本実施の形態に係るストレージ装置を含む情報処理システムの全体構成を示すブロック図である。

【図 2】 本実施の形態に係るストレージ装置における制御装置の構成を示すブロック図である。

【図 3】 本実施の形態に係るスナップショットにおけるプライマリボリューム及びセカンダリボリュームを示す図である。

【図 4】 本実施の形態に係るスナップショットにおける処理の概要を示す図である。

【図 5】 本実施の形態に係るスナップショット管理テーブルを示す図である。

【図 6】 本実施の形態に係る S - V O L 変換テーブルを示す図である。

【図 7】 本実施の形態に係るスナップショットにおける処理の概要を示す図である。

【図 8】 本実施の形態に係るスナップショットにおける処理の概要を示す図である。

【図 9】 本実施の形態に係るボリューム管理テーブルを示す図である。

【図 1 0】 本実施の形態に係る差分ボリューム管理テーブルを示す図である。

【図 1 1】 本実施の形態に係るプライマリボリュームへのデータの読み出し処理を示すフローチャートである。

【図 1 2】 本実施の形態に係るプライマリボリュームへのデータの書き込み処理を示すフローチャートである。

【図 1 3】 本実施の形態に係るセカンダリボリュームへのデータの読み出し処理を示すフローチャートである。

【図 1 4】 本実施の形態に係るセカンダリボリュームへのデータの書き込み処理を示すフローチャートである。

【図 15】 本実施の形態に係るフォーマットを説明するためのブロック図である。

【図 16】 本実施の形態に係るフォーマット管理テーブルを示す図である。

【図 17】 本実施の形態に係るフォーマットを説明するためのブロック図である。

【図 18】 本実施の形態に係るフォーマットを説明するためのブロック図である。

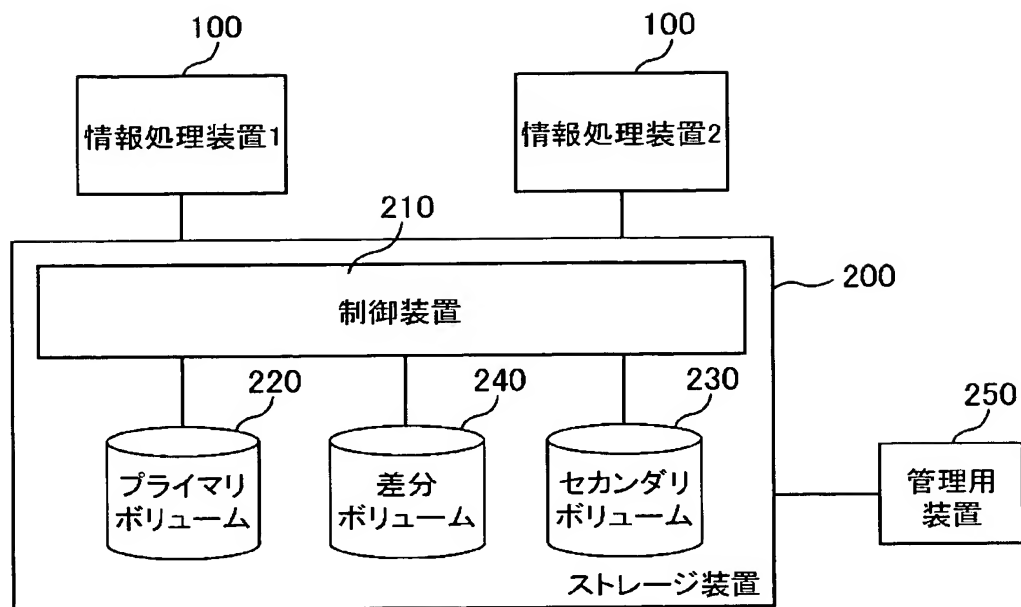
【図 19】 本実施の形態に係るフォーマット処理を示すフローチャートである。

【符号の説明】

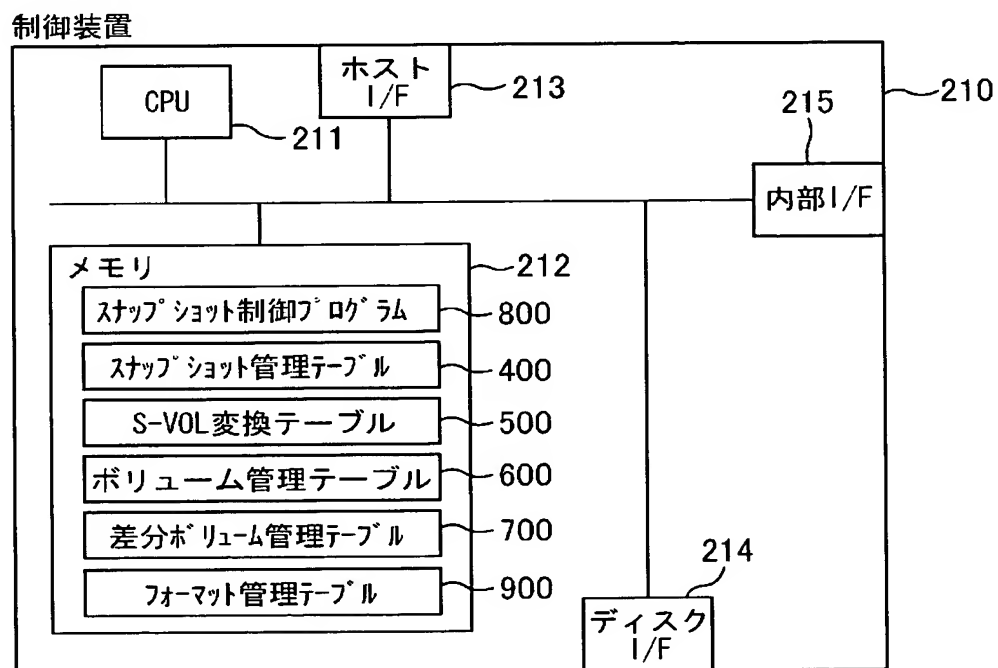
100	情報処理装置
200	ストレージ装置
210	制御装置
211	CPU
212	メモリ
220	プライマリボリューム
230	セカンダリボリューム
240	差分ボリューム
241	他の記憶ボリューム
250	管理用装置
400	スナップショット管理テーブル
500	S-VOL変換テーブル
600	ボリューム管理テーブル
700	差分ボリューム管理テーブル
800	スナップショット制御プログラム
900	フォーマット管理テーブル

【書類名】 図面

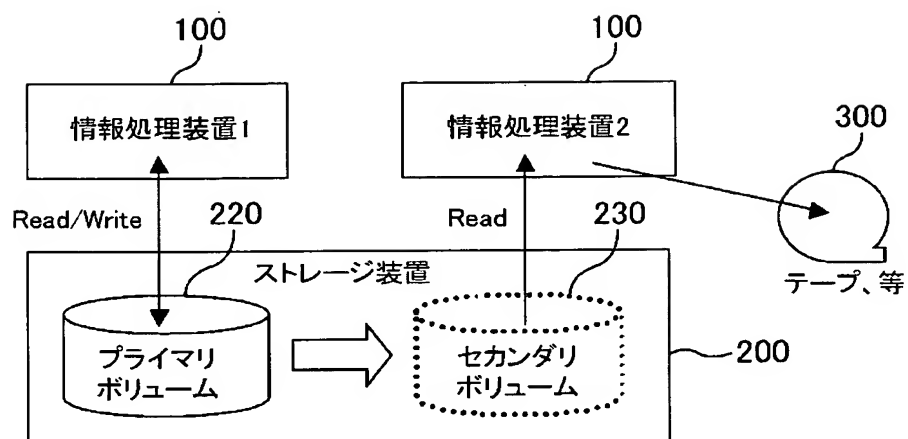
【図 1】



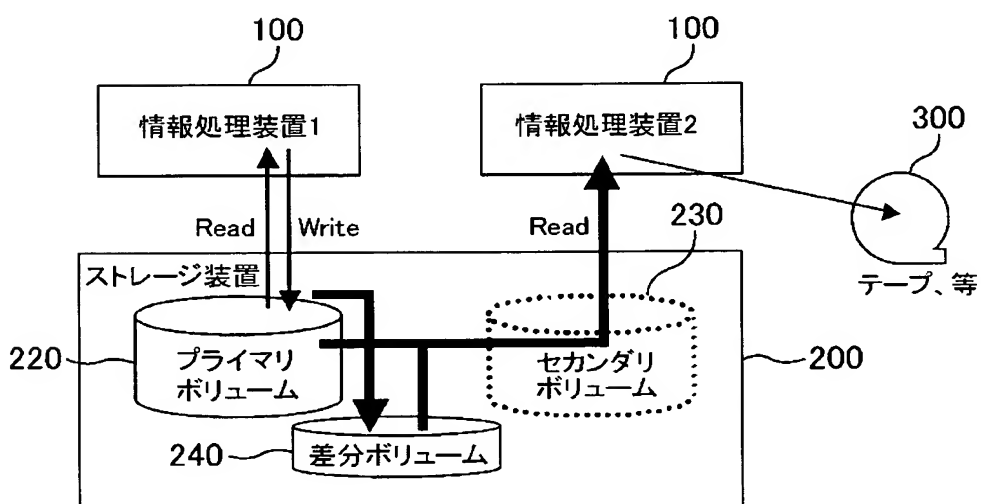
【図 2】



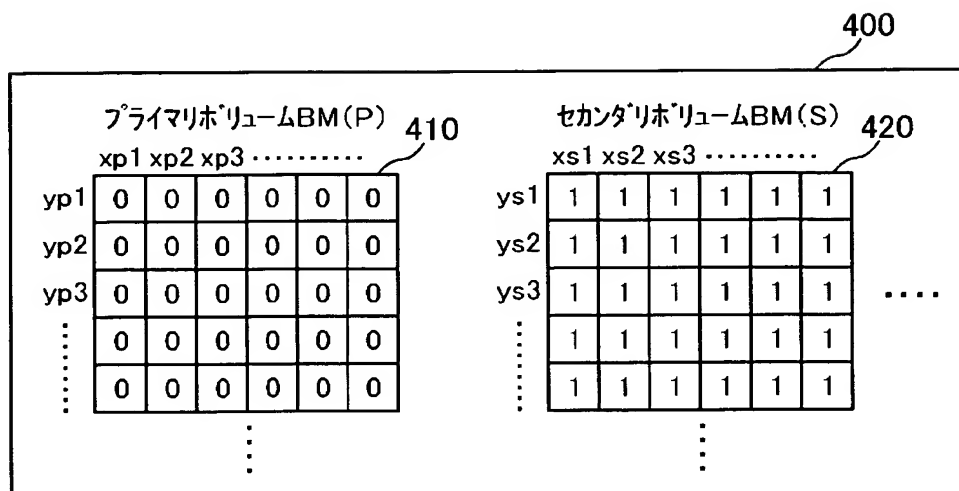
【図 3】



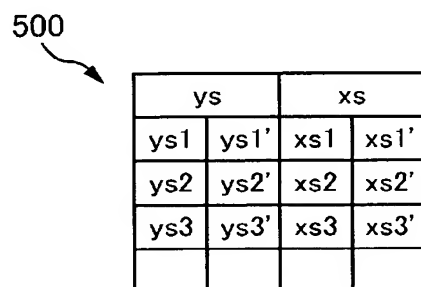
【図 4】



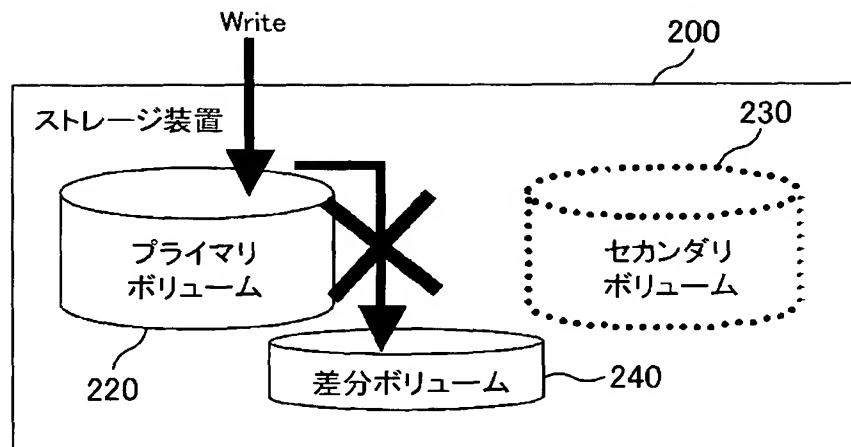
【図 5】



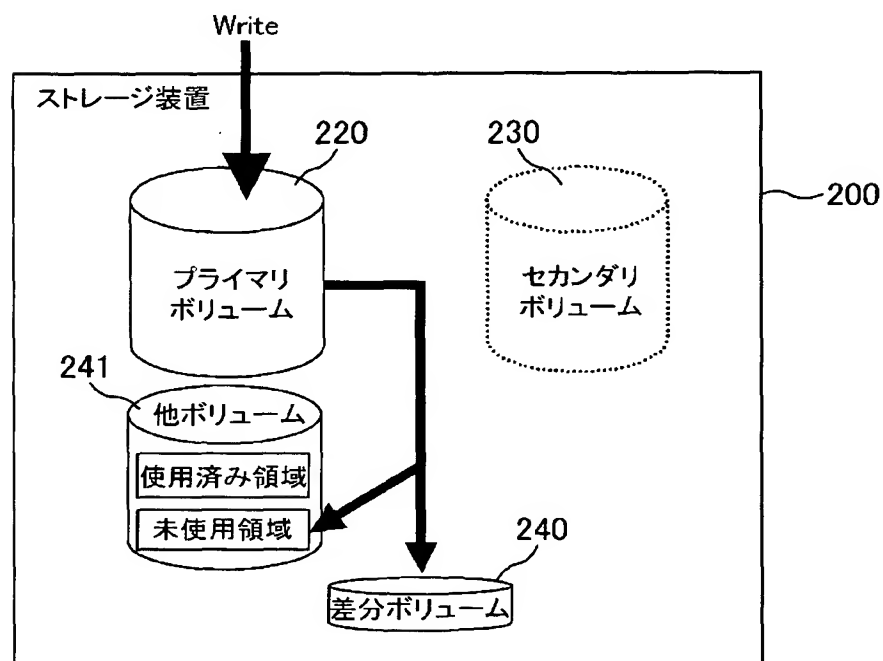
【図 6】



【図 7】

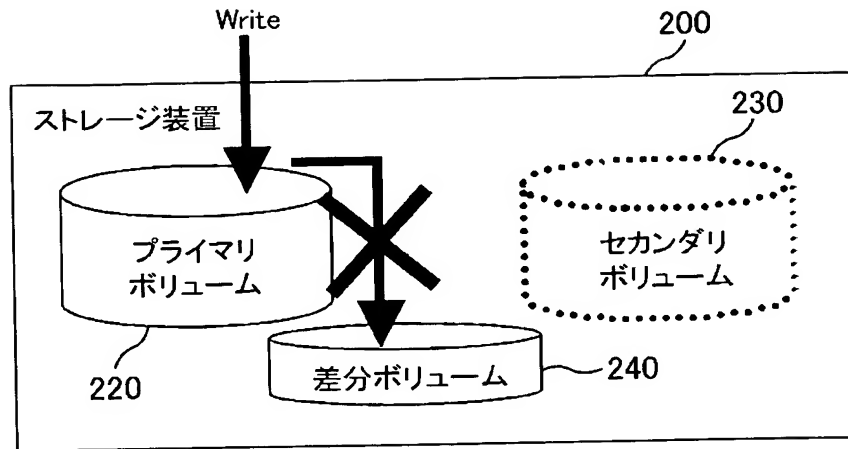


【図 8】

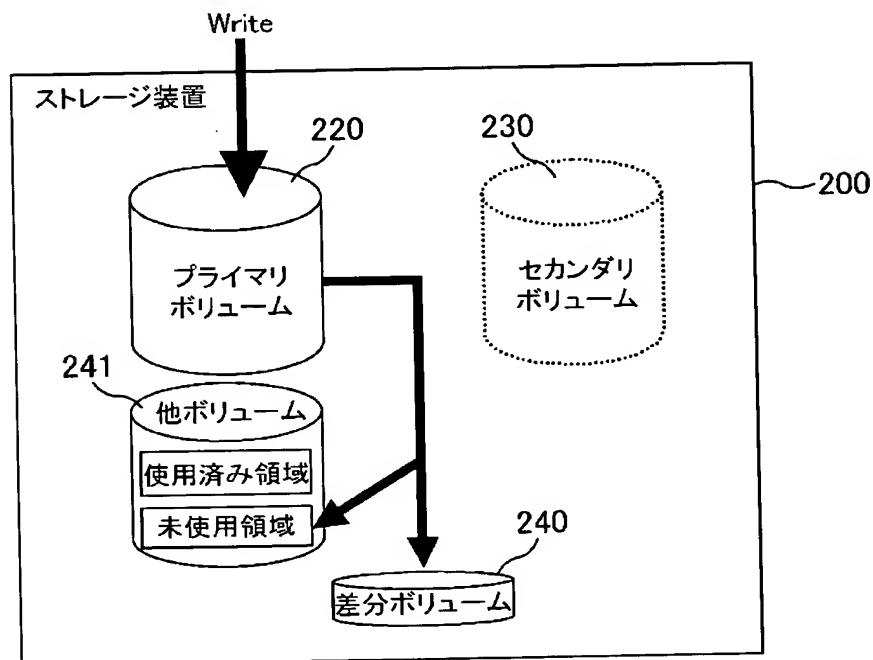


648

【図 7】



【図 8】



【図 9】

600

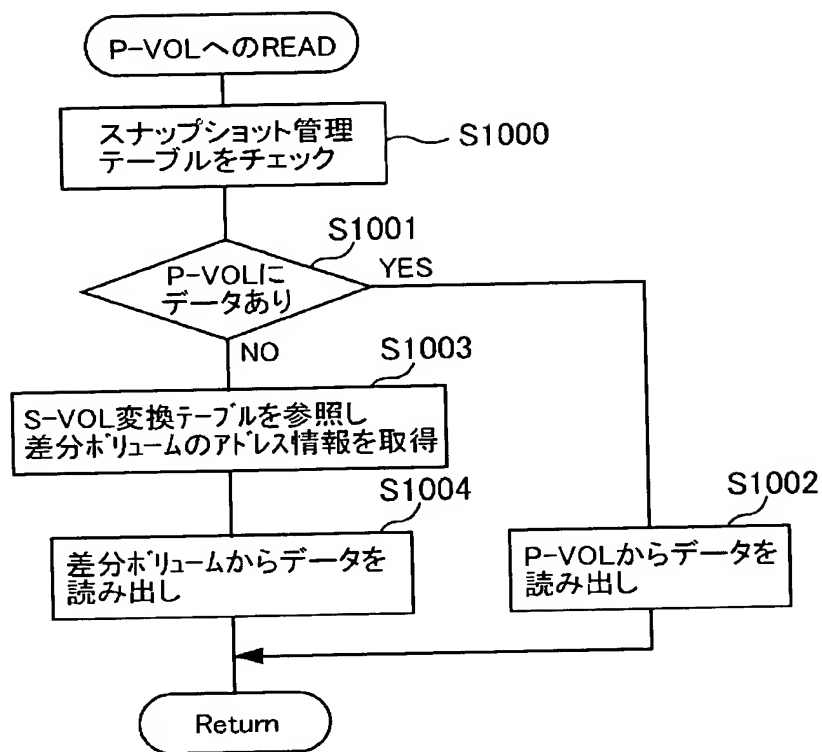
LUN	記憶 容量	空き 容量	スナップショット用ボリューム管理				用途
			ボリューム 種別	ヘア ボリューム NO	差分 ボリューム NO	差分 ボリューム 使用可否	
0	100GB	40GB	—	—	—	○	DB1
1	500GB	300GB	P	2	6	—	WEB
2	—	—	S	1	—	—	—
3	800GB	200GB	P	4	5.9	—	DB2
4	—	—	S	3	—	—	—
5	200GB	0GB	差分	—	—	○	—
6	200GB	100GB	差分	—	—	○	DB3
7	100GB	10GB	—	—	—	○	DB4
8	100GB	90GB	—	—	—	○	DB5
9	300GB	80GB	差分	—	—	○	DB6
10	300GB	200GB	—	—	—	—	FS

【図 10】

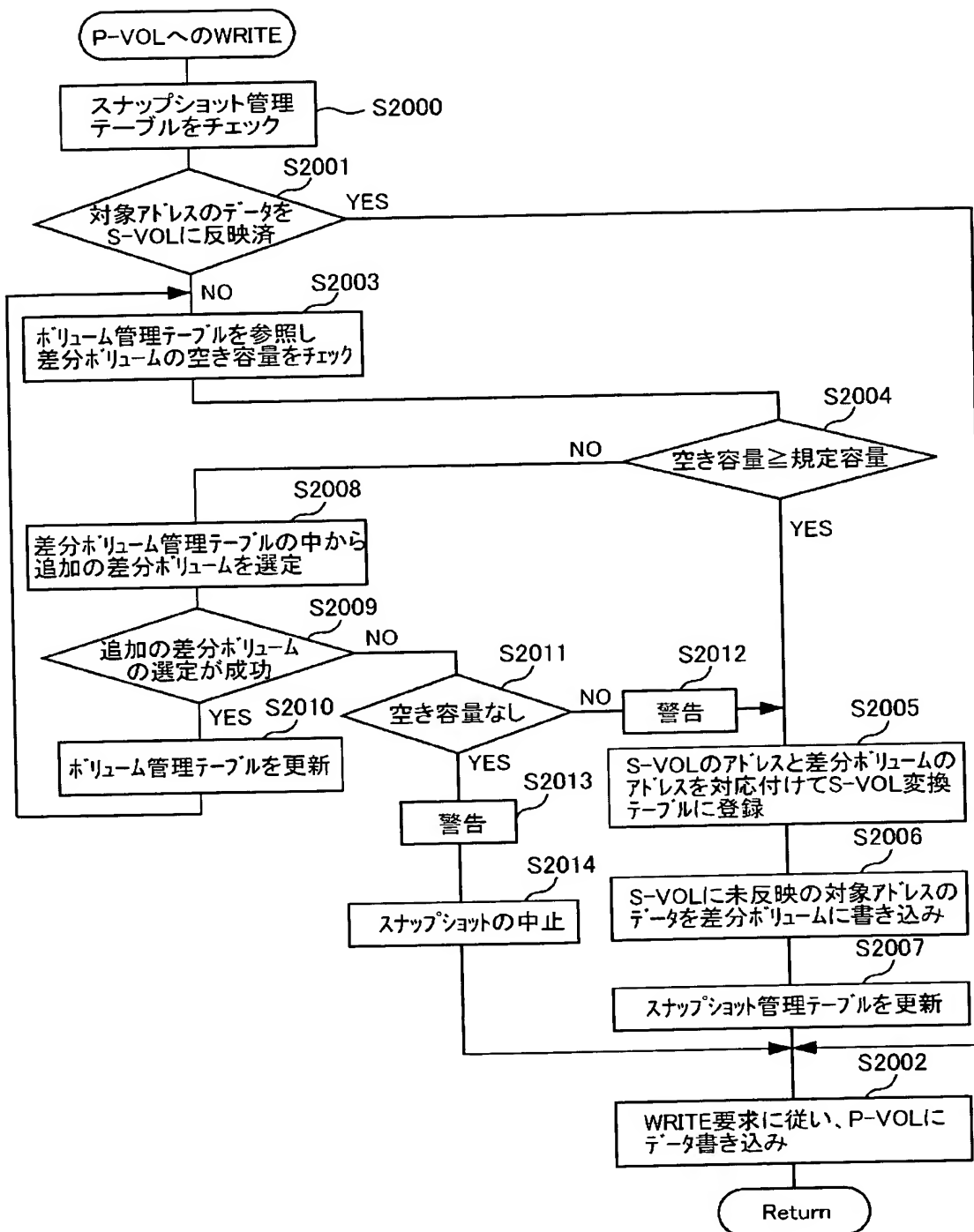
700

P-VOL NO.	差分ボリューム候補
1	LU6,LU7,LU8
3	LU5,LU9,LU0,LU7,LU8

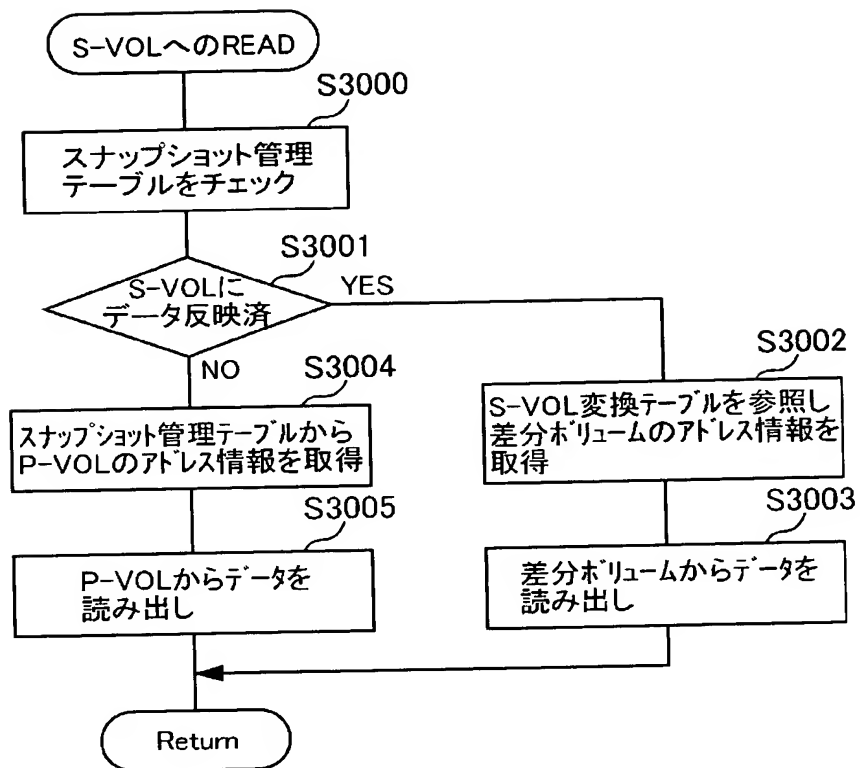
【図 11】



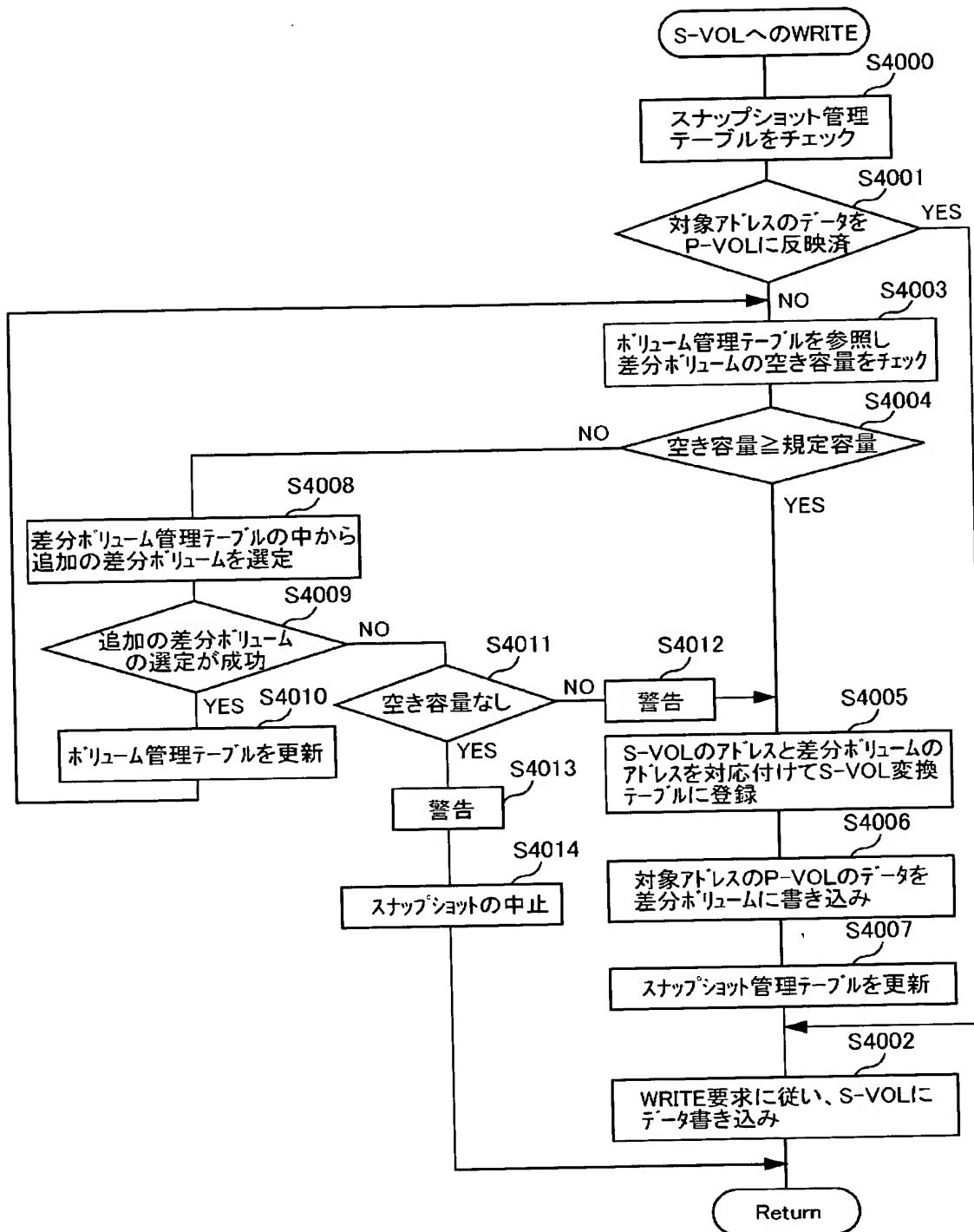
【図12】



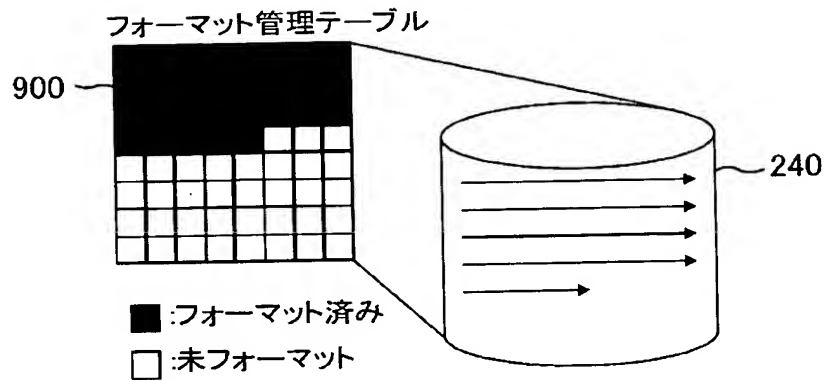
【図 13】




【図 14】



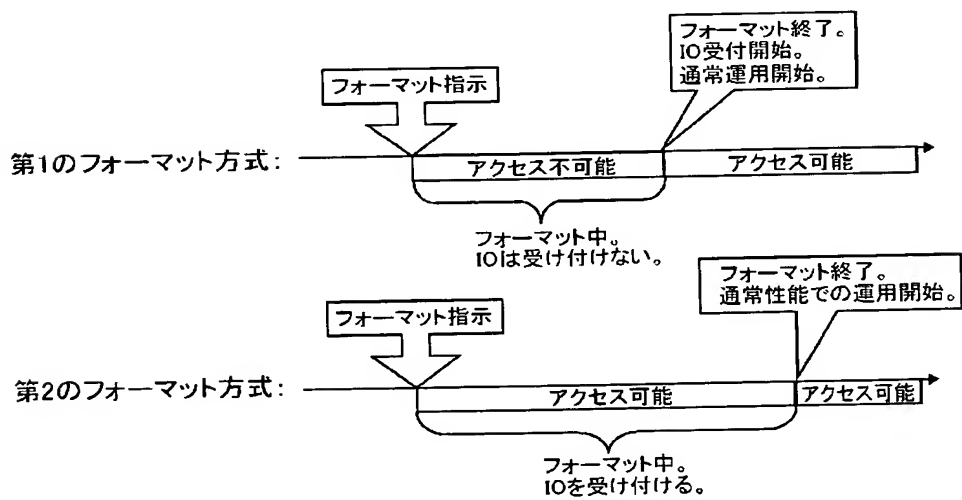
【図 15】



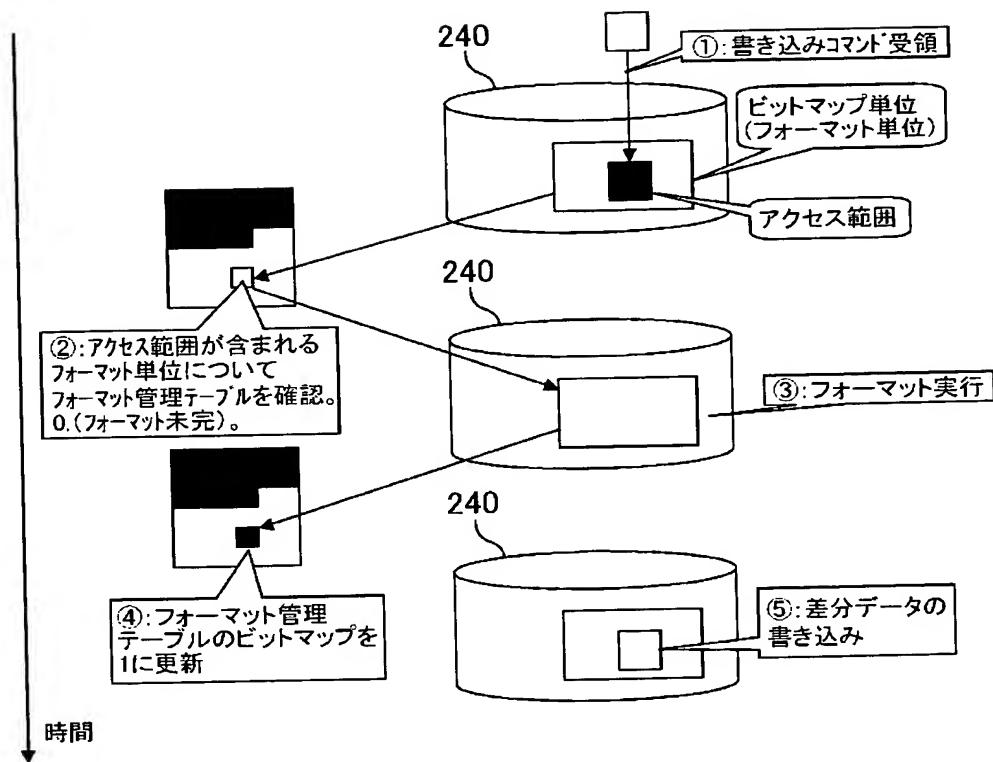
【図 16】

	xf1	xf2	xf3				
yf1	1	1	1	1	1	1	 900	
yf2	1	1	0	0	0	0		
yf3	0	0	0	0	0	0		
⋮	0	0	0	0	0	0	
⋮	0	0	0	0	0	0		
⋮	0	0	0	0	0	0		
⋮	0	0	0	0	0	0		

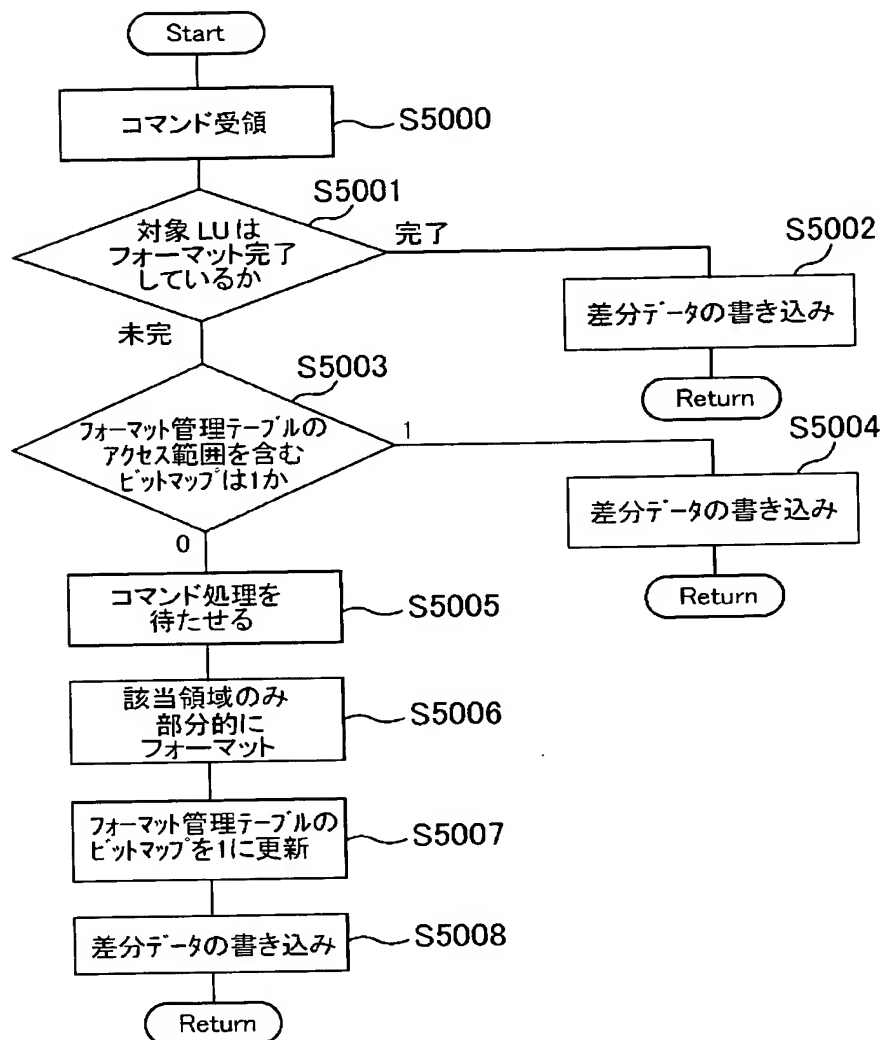
【図 17】



【図 18】



【図 19】



【書類名】 要約書

【要約】

【解決手段】 データを記憶するための複数の記憶ボリュームと、ある時点の後に、第1の前記記憶ボリュームに記憶される前記ある時点における前記データの更新要求を受信する手段と、前記更新要求の対象となった前記ある時点における前記データの複製を第2の前記記憶ボリュームに記憶する手段とを備え、前記第2の記憶ボリュームの状態に応じて、前記更新要求の対象となった前記ある時点における前記データの複製を、第3の前記記憶ボリュームの未使用の記憶領域に記憶する手段を備えるストレージ装置に関する。

【選択図】 図8